TITLE:

On a learning rate factor and extent of ordering in basic self-organizing maps (Nonlinear Analysis and Convex Analysis)

AUTHOR(S):

Hoshino, Mitsuhiro

CITATION:

Hoshino, Mitsuhiro. On a learning rate factor and extent of ordering in basic self-organizing maps (Nonlinear Analysis and Convex Analysis). 数理解析研究所講究録 2018, 2065: 147-153

ISSUE DATE:

2018-04

URL:

http://hdl.handle.net/2433/241915

RIGHT:

京都大学学術情報リポジトリ KURENAI 紅
Kyoto University Research Information Repository

数理解析研究所講究録
第2065巻 2018年 147-153

On a learning rate factor and extent of ordering in basic self-organizing maps

Mitsuhiro Hoshino

Department of Management Science and Engineering,
Faculty of Systems Science and Technology,
Akita Prefectural University,
84-4 Ebinokuchi Tsuchiya Yurihonjo, Akita 015-0055, Japan

Abstract. We deal with an essential mathematical model for self-organizing maps referred to as Kohonen type algorithm. Self-organizing map algorithm has been widely used as a useful tool in many practical problems on extensive fields. By repeating learning, some model functions in self-organizing maps have some properties of mathematical interest such as regularity between the nodes and their values. We mainly describe a learning late factor and extent of ordering in essential self-organizing maps with one-dimensionally indexed array. We present some numerical examples and make a comparison of the degree of converging for some learning processes in a basic self-organizing map.

1. Self-organizing map models and their formulation

We consider mathematical models of self-organizing maps referred to as Kohonen [8] type algorithm, which is very practical and has many useful applications, such as a semantic map, a diagnosis of speech voicing, the traveling-salesman problem, and so on. We can observe some interesting phenomena between the array of nodes and the values of nodes on iterative processes in these models. On that respect, a mathematical argument of the learning process in the one-dimensional node case with one-dimensional inputs was given by Cottrell and Fort [1]. Subsequently, convergence properties of learning processes are somewhat generally studied, e.g., in Erwin, Obermayer, and Schulten [2][3][4]. However some problems on this problem still remain without sufficient arguments on models with general conditions.

The purpose of our work is to make a study of closed classes of states and their characterization in the model. In this paper, we concentrate on a consideration of ordering of node values on learning process by several numerical examples and their calculations from various angles. One of our aim is to construct a theory on

---

Key words and phrases. self-organizing maps, learning-rate factor.
Email address: hoshino@akita-pu.ac.jp

basis of wide-ranging experiments in learning processes on basic self-organizing maps with essential structures.

We consider to characterize a model $(I, V, X, \{m_k(\cdot)\}_{k=0}^{\infty})$ with four elements which consist of the *nodes*, the *values of nodes*, *inputs* and *model functions* with some *learning processes*, in this paper. There are several types of models with various spaces of nodes, spaces of their values and ways of learning for nodes. We suppose the following in this paper.

(i) We suppose an array of *nodes*. Let I denote the set of all nodes, which is called the *node set*. We assume that I is a countable set metrized by a metric d.

(ii) We suppose that each node has its *value*. V is the space of values of nodes. We assume that V is a real linear normed space with a norm $\|\cdot\|$. A mapping $m : I \to V$ transforming each node i to its value $m(i)$ is called a *model function*. Let M be the set of all model functions.

(iii) X is the *input set*. Let X be a subset of V. $x \in X$ is called an *input*.

(iv) The *learning process* is defined by the following. If an input is given, then the value of each node is renewed to a new value by the input. If an initial model function m_0 and a sequence $x_0, x_1, x_2, \ldots \in X$ of inputs are given, then the model functions $m_1, m_2, m_3, \ldots$ are generated sequentially according to

$$m_{k+1}(i) = (1 - \alpha_{m_k,x_k}(i))m_k(i) + \alpha_{m_k,x_k}(i)x_k, \quad k = 0, 1, 2, \ldots,$$

where α_{m_k,x_k} is the learning rate which satisfies $0 \le \alpha_{m_k,x_k} \le 1$.

2. An absorbing class of a fundamental self-organizing map

In this paper, we restrict our considerations to a basic self-organizing map with real-valued nodes and a one-dimensional array of nodes. We suppose that a set V of values of nodes is identified with $\mathbb{R}$ which is the set of all real numbers.

We consider a model

$$(I = \{1, 2, \ldots, n\}, V = \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^{\infty}).$$

(i) Let $I = \{1, 2, \ldots, n\}$ be the node set with metric $d(i, j) = |i - j|$. (ii) Assume $V = \mathbb{R}$, that is, each node is $\mathbb{R}$-valued. (iii) $x_0, x_1, x_2, \ldots \in X \subset \mathbb{R}$ is an input sequence. (iv) we assume a learning process defined by the following procedures.

Learning process $\mathrm{L_A}$ with a learning radius $r = 1$ is as follows.
(a) Areas of learning:

$$I(m_k, x_k) = \{i^* \in I \mid |m_k(i^*) - x_k| = \inf_{i \in I} |m_k(i) - x_k|\} \tag{1}$$

and $N_1(i) = \{j \in I \mid |j-i| \leq 1\}$. (b) Learning-rate factor: $0 \leq \alpha \leq 1$. (c) Learning: let $N_1(I(m_k, x_k)) = \cup_{i^* \in I(m_k,x_k)} N_1(i^*)$ and $\{m_k\}$ is defined by the following, for each $k = 0, 1, 2, \ldots$, if $i \in N_1(I(m_k, x_k))$ then

$$m_{k+1}(i) = (1-\alpha)m_k(i) + \alpha x_k, \tag{2}$$

otherwise $m_{k+1}(i) = m_k(i)$.

By repeating learning, some model functions have properties such as monotonicity and a certain regularity which may appear in the relation between the array of nodes and the values of nodes. Self-organizing maps apply to many practical problems by using this property.

The following is a well-known property [8].

Theorem 1 *We consider a self-organizing map model*

$$(\{1, 2, \ldots, n\}, \mathbb{R}, X \subset \mathbb{R}, \{m_k(\cdot)\}_{k=0}^{\infty})$$

with Learning process $L_A(r = 1)$. For model functions m_1, m_2, …, the following statements hold. If m_k is increasing on I, that is $m_k(i) \leq m_k(i+1)$ for all i, then m_{k+1} is increasing on I. If m_k is decreasing on I, that is $m_k(i) \geq m_k(i+1)$ for all i, then m_{k+1} is decreasing on I. Moreover, if m_k is strictly increasing on I, that is $m_k(i) < m_k(i+1)$ for all i, then m_{k+1} is strictly increasing on I. If m_k is strictly decreasing on I, that is $m_k(i) > m_k(i+1)$ for all i, then m_{k+1} is strictly decreasing on I.

The set of states with monotone in this self-organizing map is a closed class. Such properties as monotone are called *absorbing states* of self-organizing map models.

We give a results for preserving monotone of model functions.

Theorem 2 *We consider a self-organizing map model*

$$(\{1, 2, \ldots, n\}, \mathbb{R}, X, \{m_k(\cdot)\}_{k=0}^{\infty}).$$

Assume Learning process L_A $(r = 1, 2, \ldots)$ with learning rates α_i depending on node i. For learning, let $N_r(I(m_k, x_k)) = \cup_{i^* \in I(m_k,x_k)} \{i \in I \mid |i - i^*| \leq r\}$ and suppose that

$$m_{k+1}(i) = \begin{cases} (1-\alpha_i)m_k(i) + \alpha_i x, & \textit{if } i \in N_r(I(m_k, x_k)), \\ m_k(i), & \textit{otherwise}, \end{cases}$$

where we assume that $\{\alpha_i\} \subset [0, 1)$ satisfies, for each $i^* \in I(m_k, x_k)$,

$$\alpha_i \leq \alpha_{i+1}, \qquad i = i^* - r, i^* - r + 1, \ldots, i^* - 1$$

and

$$\alpha_i \geq \alpha_{i+1}, \qquad i = i^*, i^* + 1, \ldots, i^* + r - 1.$$

Then, for model functions $m_1, m_2, \ldots$, *if* m_k *is increasing on* I, *then* m_{k+1} *is increasing on* I. *If* m_k *is decreasing on* I, *then* m_{k+1} *is decreasing on* I. *Moreover, if* m_k *is strictly increasing on* I, *then* m_{k+1} *is strictly increasing on* I. *If* m_k *is strictly decreasing on* I, *then* m_{k+1} *is strictly decreasing on* I.

A proof of Theorem 2 is in [7].

3. Numerical calculations and transition of model functions

By Theorem 2, model functions m_k turn to monotone state from non-monotone state after sufficiently many times of iterations.

Example 1

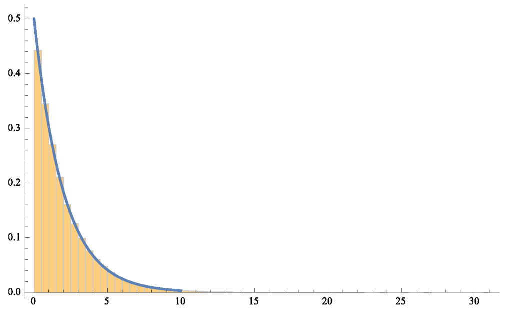

Figure 1: Histogram of input values.

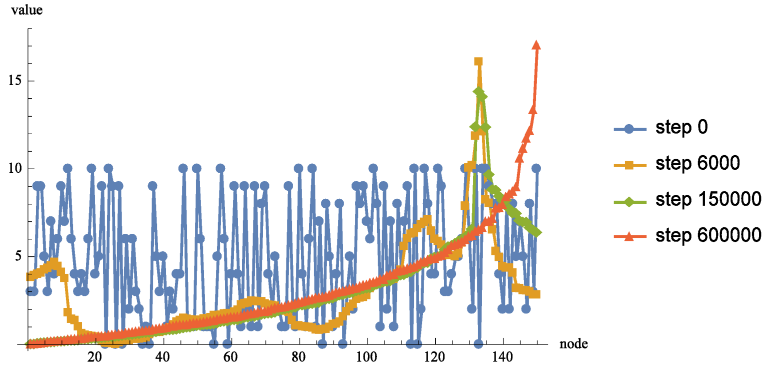

Figure 2: The horizontal and vertical axes show the node index and the value of each node, respectively.

We give some numerical examples of self-organizing map process models. Figure 2 illustrates a transition of the values of nodes in a 150 nodes model with a learning process from random inputs generated by the exponential distribution shown in Figure 1 which has mean 2 and variance 4. The initial values of nodes shown in a legend, step 0 of Figure 2, are given uniform randomly. We can observe that the values of nodes turn to ordering state gradually. The upper trajectory is Figure 3 shows a transition of the index of the node which has the maximum value. The lower one shows a transition with respect to the minimum node.

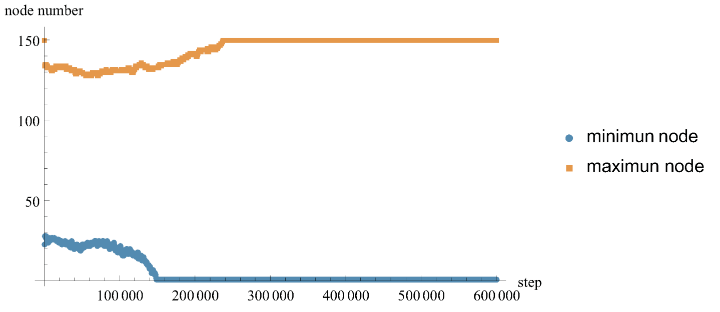

Figure 3: The horizontal and vertical axes show iteration steps and the maximum or minimum node index, respectively.

□

Example 2 On the other hand, in a learning processes on the self-organizing map model shown in Figure 4, the node values turn to ordering state gradually, however, is still not monotone state at 600,000 steps. Figure 5 shows that the node index which has the minimum value in this example does not turn to the index of either side node throughout.

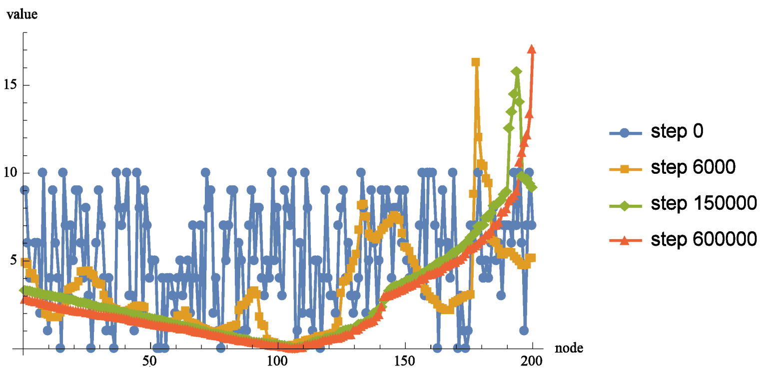

Figure 4: A transition of the node values in a 200 nodes model with the same exponential inputs as inputs in Example 1.

□

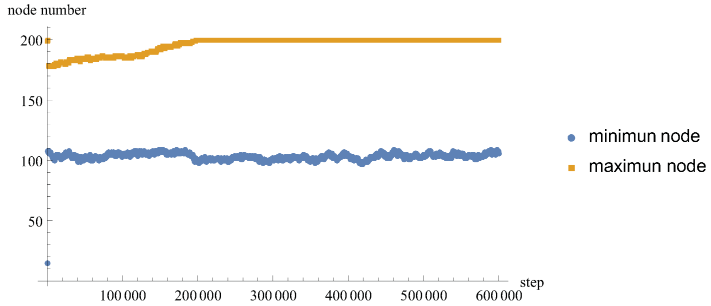

Figure 5: Transitions of the maximum and minimum node index.

Example 3 Figure 6 is given by the date which consists of the numbers of renewal times up to a monotone state by 135 experiments under 15 different initial values of nodes, 3 different input sequences from normal distribution $N(0, 2^2)$, exponential distribution with mean 0 and variance 4 and mixed distribution with probability density function $\frac{3e^{-\frac{9}{2}(x-1)^2}}{2\sqrt{2\pi}} + \frac{3e^{-\frac{9}{2}(x+1)^2}}{2\sqrt{2\pi}}$, respectively. The three horizontal lines of the rectangle indicate the mean−SD (the standard deviation), the mean, and the mean+SD, respectively. The ends of lower and upper vertical lines either side of the rectangle show the maximum and the minimum, respectively.

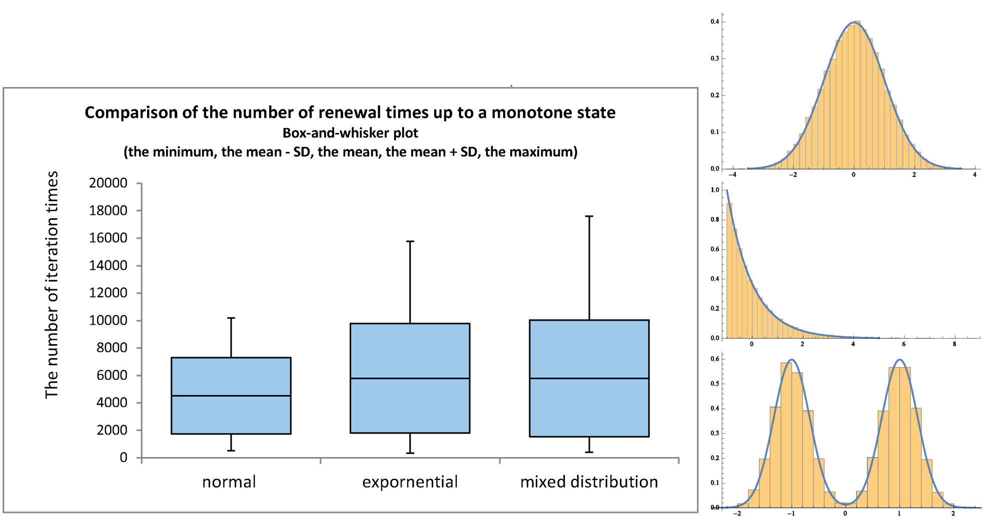

Figure 6: Box-and-whisker plot (the minimum, the mean−SD, the mean, the mean+SD, the maximum). The three pictures on the right show input distributions, a normal distribution, an exponential distribution and a mixed distribution from the top.

This figure illustrates a comparison of the number of renewal times up to a monotone state. The mean renewal times up to monotone state by normal inputs is less than the mean by exponential one, statistically, p-value of T-test is 0.0404863.

$\square$

References

[1] M. Cottrell and J.-C. Fort, *Étude d'un processus d'auto-organisation*, Annales de l'Institut Henri Poincaré, 23(1) (1987), pp. 1–20.

[2] E. Erwin, K. Obermayer and K. Schulten, *Convergence properties of self-organizing maps*, in T. Kohonen, K. Mäkisara, O. Simula and J. Kangas, editors, Artificial Neural Networks, pp. 409–414, Amsterdam Netherlands Elsevier, 1991.

[3] E. Erwin, K. Obermayer and K. Schulten, *Self-organization maps: stationary states, metastability and convergence rate*, Bio. Cybern., 67 (1992), pp. 35–45.

[4] E. Erwin, K. Obermayer and K. Schulten, *Self-organization maps: ordering, convergence properties and energy functions*, Bio. Cybern., 67 (1992), pp. 47–55.

[5] M. Hoshino and Y. Kimura, *Absorbing states and quasi-convexity in self-organizing maps*, J. Nonlinear Convex Anal., Vol. 10, No. 3 (2009), pp. 395–406.

[6] M. Hoshino and Y. Kimura, *State preserving properties in self-organizing maps with inputs in an inner product space*, Nonlinear Analysis and Convex Analysis (Tokyo, 2009), pp. 65–76.

[7] M. Hoshino, *On state preserving property and nonexpansivity in self-organizing maps*, J. Nonlinear Convex Anal. Vol. 16 Num. 11 (2015), pp. 2255–2261

[8] T. Kohonen, *Self-Organizing Maps, Third Edition*, Springer, 2001.